

Customer No. 20350
TOWNSEND and TOWNSEND and CREW LLP
Two Embarcadero Center, 8th Floor
San Francisco, California 94111-3834
(415) 576-0200

**ASSISTANT COMMISSIONER FOR PATENTS
BOX PATENT APPLICATION
Washington, D.C. 20231**

Sir:

Transmitted herewith for filing under 37 CFR 1.53(b) is the

- ☒ patent application of
☐ continuation patent application of
☐ divisional patent application of
☐ continuation-in-part patent application of

Inventor(s)/Applicant Identifier: **Manabu Kitamura, Kenji Yamagami, Akira Yamamoto, Minoru Kosuge**

For: **Computer System and Data Sharing Method Between Computers**

- ☒ This application claims priority from each of the following Application Nos./filing dates:
Japanese Patent Application No: 2000-088509, filed March 24, 2000
the disclosure(s) of which is (are) incorporated by reference.
- ☐ Please amend this application by adding the following before the first sentence: "This application is a ☐ continuation ☐ continuation-in-part of and claims the benefit of U.S. Application No. 60/_____, filed _____, the disclosure of which is incorporated by reference."

Enclosed are:

- ☒ 15 page(s) of specification
☒ 4 page(s) of claims
☒ 1 page of Abstract
☒ 9 sheet(s) of ☒ formal ☐ informal drawing(s).
☒ An assignment of the invention to Hitachi, Ltd.
☒ A ☒ signed ☐ unsigned Declaration & Power of Attorney
☐ A ☐ signed ☐ unsigned Declaration.
☐ A Power of Attorney by Assignee with Certificate Under 37 CFR Section 3.73(b).
☐ A verified statement to establish small entity status under 37 CFR 1.9 and 37 CFR 1.27 ☐ is enclosed ☐ was filed in the prior application and small entity status is still proper and desired.
☒ A certified copy of a Japanese application.
☐ Information Disclosure Statement under 37 CFR 1.97.
☐ A petition to extend time to respond in the parent application.
☐ Notification of change of ☐ power of attorney ☐ correspondence address filed in prior application.

	(Col. 1)	(Col. 2)
FOR:	NO. FILED	NO. EXTRA
BASIC FEE		
TOTAL CLAIMS	17 - 20	= *0
INDEP. CLAIMS	6 - 3	= *3
<input type="checkbox"/> MULTIPLE DEPENDENT CLAIM PRESENTED		

* If the difference in Col. 1 is less than 0, enter "0" in Col. 2.

Please charge Deposit Account No. 20-1430 as follows: \$924.00

- ☒ Any additional fees associated with this paper or during the pendency of this application.
☐ The issue fee set in 37 CFR 1.18 at or before mailing of the Notice of Allowance, pursuant to 37 CFR 1.311(b)

- ☐ A check for \$_____ is enclosed.
2 extra copies of this sheet are enclosed.

Telephone:
(415) 576-0200

Facsimile:
(415) 576-0300

Attorney Docket No. 16869P015200

"Express Mail" Label No. EL49155064US

Date of Deposit: October 25, 2000

I hereby certify that this is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated above, addressed to:

Assistant Commissioner for Patents
Washington, D.C. 20231

By: Ron Anton

COPY



SMALL ENTITY

RATE	FEE
	\$345.00
x \$9.00 =	
x \$39.00 =	
+ \$117.00 =	
TOTAL	

OR

OR

OR

OR

OR

OR

OR

OR

OR

OR

OTHER THAN SMALL ENTITY

RATE	FEE
	\$690.00
x \$18.00 =	\$0
x \$78.00 =	\$234.00
+ \$260.00 =	
TOTAL	\$924.00

Respectfully submitted,
TOWNSEND and TOWNSEND and CREW LLP

Robert C Colwell

Robert C. Colwell
Reg No.: 27,431
Attorneys for Applicant

日本国特許庁
PATENT OFFICE
JAPANESE GOVERNMENT

349900-742US/
16867P-015

J-951 U.S. PTO
09/697088
10/25/00

別紙添付の書類に記載されている事項は下記の出願書類に記載されて
いる事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed
with this Office.

出願年月日
Date of Application:

2000年 3月24日

出願番号
Application Number:

特願2000-088509

出願人
Applicant(s):

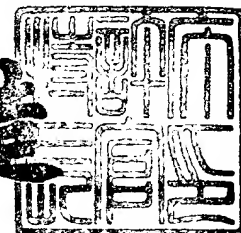
株式会社日立製作所

CERTIFIED COPY OF
PRIORITY DOCUMENT

2000年 9月29日

特許庁長官
Commissioner,
Patent Office

及川耕造



【書類名】 特許願

【整理番号】 K99007421

【提出日】 平成12年 3月24日

【あて先】 特許庁長官殿

【国際特許分類】 G06F 3/06

【請求項の数】 15

【発明者】

 【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所 システム開発研究所内

 【氏名】 北村 学

【発明者】

 【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所 システム開発研究所内

 【氏名】 山神 憲司

【発明者】

 【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所 システム開発研究所内

 【氏名】 山本 彰

【発明者】

 【住所又は居所】 神奈川県小田原市国府津 2 8 8 0 番地 株式会社日立製作所 ストレージシステム事業部内

 【氏名】 小菅 稔

【特許出願人】

 【識別番号】 000005108

 【氏名又は名称】 株式会社日立製作所

【代理人】

 【識別番号】 100075096

 【弁理士】

 【氏名又は名称】 作田 康夫

【手数料の表示】

【予納台帳番号】 013088

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 レプリケーションによるホスト間データ共有方式

【特許請求の範囲】

【請求項 1】

第 1 の計算機と、第 2 の計算機と、前記第 1 及び第 2 の計算機に接続する記憶装置サブシステムとを備える計算機システムにおけるデータの共有方法において

前記第 1 の計算機により利用されるデータを格納した第 1 の記憶装置と第 2 の記憶装置とにより、前記第 1 の記憶装置の内容と前記第 2 の記憶装置の内容とが同一となるように制御されるペア状態を形成し、

前記第 1 の記憶装置と前記第 2 の記憶装置とのペア状態を解除して、前記第 1 の記憶装置への更新が前記第 2 の記憶装置へ反映されないように制御し、

前記第 2 の計算機が利用している第 3 の記憶装置と前記第 2 の記憶装置とをスワップして、前記第 2 の計算機による前記第 3 の記憶装置へのアクセスが前記第 2 の記憶装置に対して行われるように制御することを特徴とするデータ共有方法

【請求項 2】

前記スワップ処理の後、前記第 1 の記憶装置と前記第 3 の記憶装置とでペア状態を形成することを特徴とする請求項 1 記載のデータ共有方法。

【請求項 3】

前記記憶装置サブシステムは、前記第 1 の記憶装置を有し、前記第 1 の計算機に接続する第 1 の記憶装置サブシステムと、前記第 2 及び前記第 3 の記憶装置を有し、前記第 2 の計算機に接続する第 2 の記憶装置とで構成され、前記第 1 の計算機及び前記第 1 の記憶装置サブシステムと、前記第 2 の計算機及び前記第 2 の記憶装置サブシステムが地理的に離れた場所に設置されることを特徴とする請求項 1 又は 2 記載のデータ共有方法。

【請求項 4】

第 1 の計算機と、第 2 の計算機と、前記第 1 及び第 2 の計算機に接続する記憶装置サブシステムとを備える計算機システムにおけるデータの共有方法において

前記第 1 の計算機により利用されるデータを格納した第 1 の記憶装置と第 2 の記憶装置とにより、前記第 1 の記憶装置の内容と前記第 2 の記憶装置の内容とが同一となるように制御されるペア状態を形成し、

前記第 1 の記憶装置と前記第 2 の記憶装置とのペア状態を解除して、前記第 1 の記憶装置への更新が前記第 2 の記憶装置へ反映されないように制御し、

前記第 2 の記憶装置の内容を第 3 の記憶装置へコピーし、

前記第 2 の計算機が利用している第 4 の記憶装置と前記第 3 の記憶装置とをスワップして、前記第 2 の計算機による前記第 4 の記憶装置へのアクセスが前記第 3 の記憶装置に対して行われるように制御することを特徴とするデータ共有方法。

【請求項 5】

前記第 2 の記憶装置の内容を前記第 3 の記憶装置にコピーするステップは、前記第 2 の記憶装置に格納されている前記第 1 の計算機により使用されるデータフォーマットを前記第 2 の計算機により使用されるデータフォーマットに変換するステップを含むことを特徴とする請求項 4 記載のデータ共有方法。

【請求項 6】

前記データフォーマットは、前記第 1 の計算機及び前記第 2 の計算機と前記記憶装置サブシステムとの間のインタフェースに依存するものであることを特徴とする請求項 5 記載のデータ共有方法。

【請求項 7】

前記第 1 の計算機と前記記憶装置サブシステムとの間のインタフェースがカウントキーデータ形式であって、前記第 2 の計算機と前記記憶装置サブシステムとの間のインタフェースが固定長ブロック形式であることを特徴とする請求項 6 記載のデータ共有方法。

【請求項 8】

前記第 1 乃至第 4 の記憶装置の少なくとも 1 つは、前記第 1 又は第 2 の計算機により 1 台の記憶装置として認識される論理的な記憶装置であることを特徴とする請求項 4 乃至 7 いずれかに記載のデータ共有方法。

【請求項 9】

前記第 2 の記憶装置の内容を前記第 3 の記憶装置にコピーする処理は、前記第 1 及び前記第 2 の計算機とは異なる第 3 の計算機により実施されることを特徴とする請求項 4 乃至 8 のいずれかに記載のデータ共有方法。

【請求項 10】

前記記憶装置サブシステムは、前記第 1 の記憶装置を有し、前記第 1 の計算機に接続する第 1 の記憶装置サブシステムと、前記第 2、前記第 3、及び前記第 4 の記憶装置を有し、前記第 2 の計算機に接続する第 2 の記憶装置とで構成され、前記第 1 の計算機及び前記第 1 の記憶装置サブシステムと、前記第 2 の計算機及び前記第 2 の記憶装置サブシステムが地理的に離れた場所に設置されることを特徴とする請求項 4 乃至 8 いずれかに記載のデータ共有方法。

【請求項 11】

第 1 の計算機と、第 2 の計算機と、前記第 1 及び第 2 の計算機に接続する記憶装置サブシステムとを備える計算機システムにおけるデータの共有方法において

第 1 の記憶装置のある時点における内容の写しを第 2 の記憶装置に保持しておき、

前記第 1 の計算機による前記第 1 の記憶装置の内容の更新に応じて、当該更新の内容を第 3 の記憶装置に記録し、

前記第 3 の記憶装置に記録された前記更新内容に基づいて、前記第 2 の記憶装置の内容を更新し、

前記第 2 の記憶装置を前記第 2 の計算機と接続し、前記第 2 の計算機によるデータのアクセスが、前記第 2 の記憶装置に対して行われるようにすることを特徴とするデータ共有方法。

【請求項 12】

前記第 2 の記憶装置を前記第 2 の計算機と接続するステップは、前記第 2 の計算機に接続されている第 4 の記憶装置と前記第 2 の記憶装置とをスワップすることにより行われることを特徴とする請求項 11 記載のデータ共有方法。

【請求項 1 3】

第 1 の計算機と、第 2 の計算機と、前記第 1 及び第 2 の計算機に接続し、複数の記憶装置及び該複数の記憶装置を制御する制御装置を備える記憶装置サブシステムとを有する計算機システムにおいて、前記記憶装置サブシステムは、前記第 1 の計算機から第 1 の記憶装置に書き込まれたデータを第 2 の記憶装置に重複して書き込む手段と、前記第 2 の計算機のアクセスする第 3 の論理ボリュームを前記第 2 の論理ボリュームに入れ換える手段とを有することを特徴とする計算機システム。

【請求項 1 4】

複数のインタフェースを有する記憶装置サブシステムと、前記記憶装置サブシステムへのアクセスをカウントキーデータ形式に従って行う第 1 の計算機と、前記記憶装置サブシステムへのアクセスを固定長ブロック形式に従って行う第 2、及び第 3 の計算機とを有する計算機システムにおいて、前記記憶装置サブシステムは、前記第 1 の計算機から第 1 の論理ボリュームに書き込まれたデータを第 2 の論理ボリュームへ重複して書き込む手段を備え、前記第 2 の計算機は、カウントキーデータ形式で書き込まれた前記第 2 の論理ボリュームの内容を読み出して第 3 の論理ボリュームへと書き出す手段を備え、前記第 3 の計算機は、自計算機のアクセスする第 4 の論理ボリュームを前記第 3 の論理ボリュームに入れ換える手段を備えることを特徴とする計算機システム。

【請求項 1 5】

少なくとも 1 つ以上の計算機に接続される記憶装置サブシステムにおいて、前記記憶装置サブシステムは、前記計算機からアクセスされる複数のボリュームを有し、前記計算機からの指示により前記計算機のアクセスする論理ボリュームを別の論理ボリュームに入れ換える手段を有することを特徴とする記憶装置サブシステム。

【発明の詳細な説明】

【0 0 0 1】

【発明の属する技術分野】

本発明は、情報処理システムなどにおける記憶装置システムのデータ共有方法

に係り、特に、複数のインタフェースを有する記憶装置サブシステムを備えた計算機システムにおいて、データを共有する方法に関する。

【0002】

【従来の技術】

近年、パーソナルコンピュータ（パソコン）やワークステーションなどのオープンシステムで、従来、メインフレームで行われてきた業務を行う「ダウンサイジング」が盛んに行われている。メインフレームには、従来から行われてきた業務により大量の情報が蓄積されており、メインフレームによる業務処理で蓄積されてきた情報をパソコンなどの小型のコンピュータからもアクセスしたいという要求がある。

【0003】

例えば、日経コンピュータ別冊「メインフレーム'98」（日経BP社）214ページには、企業内の計算機システムにおいて、メインフレームとオープンシステムとの複合構成の事例が紹介されている。ここに紹介される計算機システムでは、メインフレームにより売り上げ情報の処理が行われ、この結果として蓄積されたデータは、集計されて1日1回オープンシステム側のコンピュータにネットワークを経由して転送される。ユーザは、オープンシステム側のコンピュータ上でデータを参照することができる。このような形態をとることで、メインフレームで扱われる基幹業務のデータを、オープンシステムからアクセスすることが可能になる。このような方式では、メインフレームで扱われる基幹業務が一時的に停止され、全てのデータがネットワークを経由してコピーされる。扱われるデータ量が多ければ、それだけコピーの時間も多くなる。このため、ホストコンピュータやネットワークの負荷が大きくなり、さらには、長時間にわたって業務を止めなければならないといった問題が生じる恐れがある。

【0004】

このような問題の解決策の1つとして、複数のホストコンピュータから1つのディスク装置（ボリューム）にアクセスできるようにすることが考えられる。例えば、特開平9-258908号公報には、複数のインタフェースを持つ記憶装置と、その複数のインタフェースに接続されたそれぞれのホストコンピュータか

ら、そのディスク装置にアクセスする方法が開示されている。この技術によれば、システム間でのデータのコピーが不要となりホストコンピュータの負荷を軽減することができる。

【0005】

【発明が解決しようとする課題】

上述したように、特開平9-258908号公報に開示された技術によれば、複数のホストコンピュータから共通のディスク装置にアクセスすることができる。ただし、実際にデータを使用するには、そのデータが論理的に整合性のとれたものでなければならない。例えば、メインフレームからあるファイルのデータの更新が行われているときに、オープンシステムからそのファイルのデータを読み込むと、一貫性のないデータが読み込まれる恐れがある。このため、オープンシステムからのファイルデータの読み込みは、メインフレーム側で行われている処理（トランザクション）が一旦途切れて、ディスク装置に書かれたデータの内容の一貫性が保たれている間に行われる必要がある。しかし、データの内容が一貫性のある状態となるのは、あるトランザクションが完了し、次のトランザクションが始まるまでのごく短い時間でしかない。したがって、特開平09-258908号公報に開示される技術においても、メインフレーム側の処理をオフライン状態にしてデータの内容に一貫性のある状態をある程度の時間保つようにしなければ、別のホストコンピュータから共通のディスク装置に格納されているデータを読み出して利用することは事実上不可能である。

【0006】

本発明の目的は、異なるホストコンピュータ間で、アプリケーションソフトの実行を極力中断することなく、共通のファイルにアクセスできるようにすることにある。

【0007】

【課題を解決するための手段】

本発明は、上記課題を解決するために、第1の計算機と、第2の計算機と、第1及び第2の計算機に接続する記憶装置サブシステムとを備える計算機システムにおいて、第1の計算機が使用する第1の記憶装置の内容と同一の内容を持つ第

2の記憶装置を作成し、第2の記憶装置の内容を第2の計算機から使用できるようにする。

【0008】

好ましくは、第2の記憶装置は、第1の計算機による通常の処理において、第1の記憶装置に対する第1の計算機からの更新が反映されるように制御される。第2の記憶装置の内容を第2の計算機から利用できるようにする際には、第1の記憶装置への更新が第2の記憶装置に反映されないように制御される。

【0009】

本発明の一つの態様では、第2の記憶装置を、第2の計算機が使用していた第3の記憶装置と入れ替え、第2の計算機から第2の記憶装置を使用できるようにする。

【0010】

また、本発明の他の態様においては、第2の記憶装置の内容を第3の記憶装置にコピーし、第3の記憶装置を第2の計算機が使用していた第4の計算機と入れ替える。

【0011】

本発明のさらに他の態様によれば、第2の記憶装置に、ある時点における第1の記憶装置の内容の写しを保持しておく。第1の記憶装置の内容が更新されたとき、その更新の内容を第3の記憶装置に記録する。第1の記憶装置の内容を第2の計算機から使用できるようにするために、第3の記憶装置に記録された更新の内容に基づいて、第2の記憶装置の内容を書き換え、第1の記憶装置と同一の内容を第2の記憶装置に作成する。このように作成した第2の記憶装置を第2の計算機が使用していた第4の記憶装置と入れ替えることで、その内容を第2の記憶装置から使用できるようにする。

【0012】

以上のような構成を採ることにより、第1の計算機がアクセスしていたデータを第2の計算機で使うことができるようになる。

【 0 0 1 3 】

【発明の実施の形態】

ー第 1 の実施形態ー

図 1 は、本発明が適用された計算機システムの一実施形態における構成例を示す機能ブロック図である。

【 0 0 1 4 】

本実施形態における計算機システムは、ホストコンピュータ（以下、ホストと略す）1 a、ホスト 2 a、ホスト 1 a とホスト 2 a に接続される記憶装置サブシステム 3 a、及び、ネットワーク 5 を含んで構成される。

【 0 0 1 5 】

ホスト 1 a は、記憶装置サブシステム 3 a との間のインタフェース 1 1、ネットワークインタフェース 1 2、オペレーティングシステム 1 3、及び、アプリケーションプログラム 1 4 が存在する。ホスト 1 a は、物理的には、CPU、メモリなどのハードウェアを有して構成されるが、本実施形態の説明において、これらのハードウェアは、特に従来の計算機システムと変わるものではなく、本明細書では説明を省略する。このことは、本実施形態における他の装置についても同様である。オペレーティングシステム 1 3、アプリケーションプログラム 1 4 は、実際には、図示しないメモリに格納され、やはり図示されていない CPU によって実行されることで種々の機能を実現する。

【 0 0 1 6 】

ホスト 2 a は、記憶装置サブシステム 3 a との間のインタフェース 2 1、ネットワークインタフェース 2 2、オペレーティングシステム 2 3、及びアプリケーションプログラム 2 4 を有する。オペレーティングシステム 2 3、及びアプリケーションプログラム 2 4 は、オペレーティングシステム 1 3、アプリケーションプログラム 1 4 と同様に、図示しないメモリに格納され、やはり図示されていない CPU によって実行される。

【 0 0 1 7 】

記憶装置サブシステム 3 a は、複数のディスク 3 1、3 2、3 3、ホスト 1 a と接続するインタフェース 4 1、ホスト 2 a と接続するインタフェース 4 2、デ

ディスク制御装置 4 3、及びコンソール 4 4 を有する。ディスク 3 1、3 2、3 3 は、それぞれ単体の磁気ディスク装置、あるいは、ディスクアレイのように、複数の物理的な磁気ディスク装置（ドライブ）をまとめて、ホストから見て 1 または複数の論理的なディスク装置（論理ボリューム）として見えるものでも良い。後者の場合、ディスク 3 1、3 2、3 3 は、それぞれ論理ボリュームである。本実施形態では、ホスト 1 a からはディスク 3 1 に、またホスト 2 a からはディスク 3 2、3 3 にアクセスできるものとする。

【0 0 1 8】

ディスク制御装置 4 3 は、その機能として、ディスクのミラーリング機能と、ディスクマッピング変更機能を有する。

【0 0 1 9】

ミラーリング機能は、ある時点のディスクの内容をそのまま別ディスクにコピーとして残すことのできる機能である。ミラーリング機能の具体的な例として、ディスク 3 1 のコピーをディスク 3 2 にとる場合について考える。ここでは、コピー元となるディスク 3 1 を正ディスク、コピー先となるディスク 3 2 を副ディスクと呼ぶ。

【0 0 2 0】

ユーザは、記憶装置サブシステム 3 a に対して、ディスク 3 1、3 2 をそれぞれ正／副ディスクとして関連付け、コピー開始を指示する。この処理をペア形成と呼ぶ。ここでの指示は、例えば、ホスト 1 a からインタフェース 1 1 を介して、あるいは、コンソール 4 4 から出される。コンソール 4 4 から指示を与える場合には、ユーザが直接コンソール 4 4 を操作して指示を出す他、ホスト 1 a からネットワーク経由でコンソール 4 4 に指示を出し、これを受けてコンソール 4 4 が記憶装置サブシステム 3 a に指示を出すようにしてもよい。本実施形態ではホスト 1 a からネットワーク経由でコンソール 4 4 に対して指示が出され、コンソールからの指示に基づいてコピーが開始されるものとして説明するが、いずれの方法をとっても本発明のおける機能は実現できる。

【0 0 2 1】

ペア形成が指示されると、ディスク制御装置 4 3 がディスク 3 1 からディスク

3 2 へのコピーを実施する。コピー開始時点では、ディスク 3 1、3 2 のデータ内容は一致していない。この状態を「ペア形成中」と呼ぶ。ペア形成中状態では、記憶装置サブシステム 3 a は、ディスク 3 1 からディスク 3 2 へのデータのコピーを実施する。ディスク制御装置 4 3 は、この間も、ホスト 1 a からディスク 3 1 へのリード/ライトアクセスがあれば、これを受け付ける。ホスト 1 a からライトアクセスがあつてディスク 3 1 の内容を更新したとき、ディスク制御装置 4 3 は、その更新内容をディスク 3 2 にも反映する。

【 0 0 2 2 】

コピー処理が実施されて、ディスク 3 1 の内容がディスク 3 2 に完全にコピーされた状態を「duplex状態」と呼ぶ。duplex状態のとき、ホスト 1 a からディスク 3 1 へライトアクセスがあると、ディスク制御装置 4 3 は、その内容をディスク 3 2 へも反映させる。

【 0 0 2 3 】

記憶装置サブシステム 3 a がduplex状態で動作中、ディスク 3 1 のある時点の状態をディスク 3 2 に保存したい時、ユーザは、記憶装置サブシステム 3 a に対して「ペア解除」の指示を送る。記憶装置サブシステム 3 は、ペア解除の指示を受けると、ディスク 3 2 をその時点の状態で保存する。すなわち、「ペア解除」の指示を受けると、ディスク制御装置 4 3 は、これ以降、ホスト 1 a からディスク 3 1 へライトアクセスによりディスク 3 1 の内容を更新しても、ディスク 3 2 へはその更新を反映しない。この状態を「simplex状態」と呼ぶ。

【 0 0 2 4 】

simplex状態になった後、再度ディスク 3 1 とディスク 3 2 をduplex状態にしたい時、ユーザは、記憶装置サブシステム 3 a に対して、ディスク 3 1、3 2 間でのコピー開始を指示する。この場合、ディスク 3 1 とディスク 3 2 の多くのデータは同じ内容であり、コピー量は少なく、比較的短時間でduplex状態へと移行できる。

【 0 0 2 5 】

ミラーリング機能では、あるディスクに対してもう 1 台ディスクを用意してペアを形成しておき、duplex状態からsimplex状態に変えることで、任意時点でのデ

ィスクのコピーを作成することができる。

【 0 0 2 6 】

例えば、ディスクが S C S I (Small Computer System Interface) ディスクの場合、各ディスクにはターゲット I D、L U N (Logical Unit Number) が付与されている。各ホストは、これらの識別子によりインタフェースから見える複数のディスクを識別し、アクセス対象のディスクを選定する。記憶装置サブシステム 3 a は、ディスクマッピング変更機能により、ホスト 1 a、2 a に認識させるこれらの識別子を自由に変更できる。例えば、最初、ディスク 3 2 はターゲット I D が 0、L U N が 0 であり、ディスク 3 3 はターゲット I D が 1、L U N が 0 である場合に、記憶装置サブシステム 3 a は、ある時点でこれを逆にすることができる。すなわち、ディスク 3 2 のターゲット I D が 1、L U N が 0 であり、ディスク 3 3 のターゲット I D が 0、L U N が 0 であるように切り替えることができる。

【 0 0 2 7 】

あるいは、ディスク 3 3 には最初ターゲット I D、L U N を付与せず、ホストからは見えないようにしておき、ある時点でディスク 3 3 にターゲット I D とし 1、L U N とし 0 を付与し、ホストから見えるようにすることもできる。各ホストはターゲット I D、L U N でアクセス対象とするディスクを識別している。したがって、マッピング変更機能を用いることで、ホストに意識させることなく、任意の時点で、アクセス先のディスクを切り替えることが可能となる。

【 0 0 2 8 】

図 2 は、ディスク 3 1 に格納されたホスト 1 a のデータをホスト 2 a で使用するときに行われる処理の流れを示すフローチャートである。ここでは、以下の点を想定して説明を行う。

【 0 0 2 9 】

ディスク 3 1 とディスク 3 2 が duplex 状態にされており、ホスト 1 a からは少なくともディスク 3 1 が、ホスト 2 a からは少なくともディスク 3 3 が見えるように制御されているものとする。ディスク 3 2 は、ホスト 1 a、2 a から見えていても見えていなくても良いが、ここではホスト 1 a、2 a からは見えない状態

にあるものとする。

【 0 0 3 0 】

ホスト 1 a 上のアプリケーションプログラム 1 4 は、データベースなどの O L T P (OnLine Transaction Processing) 処理を行い、ディスク 3 1 への更新を頻繁に行うプログラム、ホスト 2 のアプリケーションプログラム 2 4 はディスク 3 3 へのライトアクセスは行われないプログラムである。アプリケーションプログラム 2 4 が扱えるデータのフォーマットは、アプリケーションプログラム 1 4 と同じであり、最初の状態において、ディスク 3 3 にはディスク 3 1 のある時点でのデータのコピーが格納されている。アプリケーションプログラム 1 4 と、アプリケーションプログラム 2 4 には、ネットワーク 5 を介したアプリケーション間通信機構 1 4 1、2 4 1、及びインタフェース 1 1、2 1 を介してディスクのミラーリングの機能やマッピング変更の機能の実行を指示するディスク制御機構 1 4 2、2 4 2 が含まれる。

【 0 0 3 1 】

以上のような想定の下で、本実施形態における処理は、以下のように実施される。

【 0 0 3 2 】

アプリケーションプログラム 1 4 は、まず、現在実行されているトランザクション処理を完了させる(ステップ 1 0 0 1)。これにより、ディスク 3 1 の内容が無矛盾な状態となる。次に、ディスク制御機構 1 4 2 を使用して記憶装置サブシステム 3 にペア解除を指示し、ディスク 3 1、3 2 を simplex 状態にする(ステップ 1 0 0 2)。ペア解除の後、アプリケーションプログラム 1 4 は、トランザクション処理を再開する(ステップ 1 0 0 3)。

【 0 0 3 3 】

ステップ 1 0 0 4 でアプリケーションプログラム 1 4 は、アプリケーション間通信機構 1 4 1 を使用し、ホスト 2 a に対してアプリケーションプログラム 1 4 での処理、すなわち、ディスクの simplex 状態への移行が完了したことを通知する。アプリケーションプログラム 2 4 は、この通知を受けて、ディスク 3 3 へのアクセス処理を一旦終了する(ステップ 1 1 0 1)。それから、アプリケーシ

ンプログラム 2 4 は、ディスク制御機構 2 4 2 を使用して、ディスク 3 2、ディスク 3 3 とのマッピングを変更する（ステップ 1 1 0 2）。アプリケーションプログラム 2 4 のディスクアクセス処理は、必ずしも終了する必要はなく、例えば、データベースなどの検索処理を行っている場合、1 つの処理が完了した時点であれば良い。

【 0 0 3 4 】

ステップ 1 1 0 3 で、アプリケーションプログラム 2 4 は処理を再開し、ステップ 1 1 0 4 では、アプリケーション間通信機構 2 4 1 を使用して、ホスト 1 にアプリケーションプログラム 2 4 での処理が完了したことを通知する。

【 0 0 3 5 】

アプリケーションプログラム 2 4 からの通知を受けると、アプリケーションプログラム 1 4 は、再びディスク 3 1、3 2 をペア状態にする（ステップ 1 0 0 5）。

【 0 0 3 6 】

以上の処理により、アプリケーションプログラム 1 4 により使用されていたディスク 3 1 のデータをアプリケーションプログラム 2 4 で使用できるようになる。この処理を繰り返すことにより、アプリケーションプログラム 2 4 は、アプリケーションプログラム 1 4 が使用する最新のデータを使用できる。

【 0 0 3 7 】

本実施形態では、ステップ 1 1 0 2 でのディスク 3 2、3 3 のマッピング変更は、記憶装置サブシステム 3 a のディスク制御装置 4 3 が行うが、記憶装置サブシステム 3 a がマッピング変更機能を持たない場合には、あらかじめホスト 2 a からディスク 3 2、3 3 を見えるようにしておき、ホスト 2 a 側で、ディスク 3 3 をアンマウント、その後ディスク 3 2 をディスク 3 3 がマウントされていたディレクトリに対してマウントすることで、マッピング変更と同じ処理を実現できる。

【 0 0 3 8 】

このように 2 つのシステムで同一のデータを使用するアプリケーションとしては、データマイニング、OLAP (Online Analytical Processing) などが挙げら

れる。

【 0 0 3 9 】

本実施形態によれば、例えば、ホスト 1 a において、アプリケーションプログラム 1 4 として、トランザクション型のデータベース処理を行う DBMS を適用して会計処理などを行い、ホスト 2 a において、アプリケーションプログラム 2 4 として、OLAP サーバプログラムを動作させ、ホスト 1 での処理内容を利用するような場合、ホスト 1 a の DBMS による会計処理をほとんど止めることなく、ホスト 2 a の OLAP サーバプログラムで会計処理の結果であるデータを利用することが可能となる。

【 0 0 4 0 】

－ 第 2 の実施形態 －

図 3 は、本発明の第 2 の実施形態における計算機システムの構成を示すブロック図である。

【 0 0 4 1 】

本実施形態の計算機システムは、ホスト 1 b、ホスト 2 b、ホスト 1 b とホスト 2 b に接続される記憶装置サブシステム 3 b、及びホスト 1 b と 2 b とを接続するネットワーク 5 を含んで構成される。

【 0 0 4 2 】

ホスト 1 b は、記憶装置サブシステム 3 b との間のインタフェース 1 1、ネットワークインタフェース 1 2、オペレーティングシステム 1 3、及びアプリケーションプログラム 1 6 を備える。

【 0 0 4 3 】

ホスト 2 b は、記憶装置サブシステム 3 b との間のインタフェース 2 1、ネットワークインタフェース 2 2 を有し、オペレーティングシステム 2 3、アプリケーションプログラム 2 6、及び変換プログラム 2 5 を備える。

【 0 0 4 4 】

記憶装置サブシステム 3 b は、ディスク 3 1 ～ 3 3 に加え、さらに、第 4 のディスク 3 4 を備えており、第 1 の実施形態における記憶装置サブシステム 3 a と同様の機能を有する。

【 0 0 4 5 】

本実施形態では、ホスト 1 b のアプリケーションプログラム 1 6 が扱うデータのフォーマットと、ホスト 2 b のアプリケーションプログラム 2 6 が扱うデータのフォーマットは異なる。変換プログラム 2 5 は、アプリケーションプログラム 1 6 により扱われるデータのフォーマットをアプリケーション 2 6 が扱うデータのフォーマットに変換するプログラムである。

【 0 0 4 6 】

本実施形態では、ディスク 3 1 のデータをホスト 2 b で使用できるようにするために、ディスク 3 1、3 2 は duplex 状態にしておく。ディスク制御装置 4 3 は、ホスト 1 b からは少なくともディスク 3 1 が見え、ホスト 2 b からはディスク 3 2、3 3、3 4 が見えるよう記憶装置サブシステム 3 b を制御する。アプリケーションプログラム 2 6 は、ディスク 3 4 を使用して処理を行う。最初の時点でディスク 3 4 には、変換プログラム 2 5 により変換されたディスク 3 1 のある時点のデータが格納されている。

【 0 0 4 7 】

図 4 は、ディスク 3 1 のデータをホスト b 側で処理できるようにする際に行われる処理のフローチャートである。

【 0 0 4 8 】

本処理において、ホスト 1 b により行われる処理は、第 1 の実施形態におけるホスト 1 a の処理と同じである。ステップ 1 0 0 1 から 1 0 0 4 で、ディスク 3 1 と 3 2 のペア状態を解除し、ディスク 3 2 をホスト 2 b から利用できる状態とする。

【 0 0 4 9 】

ホスト 2 b は、ホスト 1 のアプリケーションプログラム 1 6 から通知を受けると変換プログラム 2 5 によりディスク 3 2 の内容を読み出し、アプリケーションプログラム 2 6 によりアクセスできる形式に変換してディスク 3 3 に格納する（ステップ 1 2 0 1）。アプリケーションプログラム 2 6 は、ディスク 3 4 に格納されているデータを使用して処理を行うため、変換プログラム 2 5 の処理中も通常の処理を継続することが可能である。

【0050】

変換プログラム25による変換処理が完了すると、ホスト2bは、アプリケーションプログラム26によるディスクアクセス処理を一旦終了する（ステップ1101）。そして、ディスク33、ディスク34とのマッピングを変更し、アプリケーションプログラム26がデータアクセスを行うディスクをディスク34からディスク33に変更する（ステップ1102b）。

【0051】

ディスク33と34をスワップした後、アプリケーションプログラム26によるディスクアクセス処理を再開し（ステップ1103）、ホスト1bにアプリケーションプログラム26によるディスクのスワップ処理が完了したことを通知する（ステップ1104）。

【0052】

アプリケーションプログラム16は、ホスト2bからの通知を受けて、再びディスク31、32をペア状態にする。以上の処理により、アプリケーションプログラム16で使用していたディスク31と同じ内容のデータをアプリケーションプログラム26が使用することができるようになる（ステップ1005）。

【0053】

－第3の実施形態－

図5は、本発明の第3の実施形態における計算機システムの構成例を示すブロック図である。

【0054】

本実施形態の計算機システムは、ホスト1c、ホスト2c、そしてホスト1bとホスト2cに接続される記憶装置サブシステム3c、各ホストを相互に接続するネットワーク5、及び、記憶装置サブシステム3cに付随したプロセッサ6を含んで構成される。

【0055】

ホスト1cは、記憶装置サブシステム3cとの間のインタフェースとして、チャネルインタフェースのような、カウントキーデータ形式（以下、CKD形式と呼ぶ）に従うデータを扱う可変長インタフェース11、ネットワークインタフェ

ース12、オペレーティングシステム13、及びアプリケーションプログラム16を有する。

【0056】

ホスト2cは、記憶装置サブシステム3cとの間のインタフェースとして例えば、SCSIインタフェースのような固定長データ形式を扱う固定長インタフェース21、ネットワークインタフェース22、オペレーティングシステム23、及びアプリケーションプログラム26を有する。

【0057】

記憶装置サブシステム3cは、複数のディスク31、32、33、34、ホスト1cと接続する可変長インタフェース41、ホスト2cと接続する固定長インタフェース42、ディスク制御装置43、プロセッサ6と接続する固定長インタフェース45を含んで構成される。ディスク制御装置43は、ホスト1cからは、少なくともディスク31に、ホスト2cからは少なくともディスク34に、プロセッサ6からは少なくともディスク32、33にアクセスできるよう記憶装置サブシステム3cを制御する。

【0058】

プロセッサ6は、記憶装置サブシステム3cとの間の固定長インタフェース61、ネットワークインタフェース62を有する。プロセッサ6は、また、そのメモリ上に、オペレーティングシステム63、変換プログラム64を保持している。

【0059】

本実施形態では、第2の実施の形態と同様、アプリケーションプログラム26が扱えるデータフォーマットは、アプリケーション16のものとは異なる。さらに、ホスト1cは、CKD形式に従ってディスク31にデータを格納する。プロセッサ6の変換プログラム64は、CKD形式に従って記憶装置サブシステム3cに格納されたデータを固定長形式に従った形で読み出し、それをアプリケーションプログラム16により扱われるデータの形式からアプリケーションプログラム26により扱われるデータの形式に変換する。

【 0 0 6 0 】

変換プログラム 6 4 が記憶装置サブシステム 3 c から C K D 形式に従って格納されたデータを読み出すには、例えば、特開平 9 - 2 5 8 9 0 8 号公報に開示された技術を適用することができる。すなわち、変換プログラム 6 4 は、ホスト 1 c 上で動作するアプリケーションプログラム 1 6 からデータの格納位置を表す情報をネットワーク 5 を介して受け取る。そして、その情報に基づき、固定長形式での格納位置を求めて記憶装置サブシステム 3 c からデータを読み出す。

【 0 0 6 1 】

図 6 は、ディスク 3 1 のデータをホスト 2 c で使用できるようにする処理のフローチャートである。

【 0 0 6 2 】

本実施形態においても、処理の開始前の状態として、ディスク 3 1、3 2 は基本的に duplex 状態にしておく。ディスク制御装置 4 3 は、ホスト 1 c からディスク 3 1 が、ホスト 2 c からディスク 3 4 が、プロセッサ 6 からはディスク 3 2、3 3、3 4 が見えるように記憶装置サブシステム 3 c を制御する。

【 0 0 6 3 】

本実施形態において、ステップ 1 0 0 1 ~ 1 0 0 4 においてホスト 1 c により第 1 あるいは第 2 の実施形態におけるステップ 1 0 0 1 ~ 1 0 0 4 と同様の処理が行われる。なお、ホスト 1 c からプロセッサ 6 へのデータの格納場所に関する情報の転送は、この処理の開始前に予め通知しておき、プロセッサ 6 で保持しておくか、あるいは、ステップ 1 0 0 4 においてホスト 1 c 側でのペア状態の解除処理完了を通知の際に転送するようにすることができる。

【 0 0 6 4 】

変換プログラム 6 4 は、ホスト 1 c からの通知を受けて、ディスク 3 2 の内容を読み出し、アプリケーションプログラム 2 6 からアクセスできる形式に変換してディスク 3 3 に書き込む。この処理の間、アプリケーションプログラム 1 6 は、ディスク 3 1 を使用して通常の処理を継続することができる。同様に、アプリケーションプログラム 2 6 は、ディスク 3 4 を使用して、通常の処理を行うことが可能である（ステップ 1 3 0 1）。

【0065】

ディスク32からディスク33へのデータの移動（コピー）を終えると、変換プログラム64は、アプリケーション間通信機構641を使用しホスト2cに処理の完了を通知する（ステップ1302）。

【0066】

アプリケーションプログラム26は、プロセッサ6からの通知を受けてディスク34へのアクセス処理を一旦終了する（ステップ1101）。ディスク34へのアクセス処理を停止後、アプリケーションプログラム26は、ディスク制御機構242を使用して、ディスク33とディスク34とのマッピングを変更し、ディスク33とディスク34を入れ替える（ステップ1102b）。その後、アプリケーションプログラム26は、入れ替え後のディスク34へのアクセス処理を再開し（ステップ1103）、ホスト1cにアプリケーションプログラム26での処理が完了したことを通知する（ステップ1104）。

【0067】

アプリケーションプログラム16は、ホスト2cからの通知を受けると再びディスク31、32をペア状態にする（ステップ1005）。

【0068】

以上の処理により、インタフェース形式の異なるホスト間でデータを共用することが可能となる。

【0069】

－第4の実施形態－

以下に説明する第4の実施形態における計算機システムは、第3の実施形態における計算機システムと同様の構成を有する。ただし、本実施形態では、以下の点において第3の実施形態と相違する。

【0070】

ホスト1cのアプリケーションプログラム16は、通常の処理として、ディスク31に対するデータの書き込み、更新を行い、そのアドレスと更新データをディスク32に記録する。つまり、本実施形態では、ディスク32は、ログディスクとして用いられ、そのデータは、障害などの要因でアプリケーションプログラ

ム 1 6 の処理が中断された場合のデータ復旧に使用される。

【 0 0 7 1 】

更新データは、ディスク 3 2 の先頭から順に記録される。また一連の更新データは、トランザクション毎に管理され、各トランザクションごとに「チェックポイント」と呼ばれる識別 ID が記録される。アプリケーションプログラム 1 6 のディスク 3 1 へのデータ書き込みがあるトランザクションの途中で中断した場合、ログディスクに記録された最新のチェックポイントの直前までの更新履歴を参照してデータの修正を行うことで、そのチェックポイントに対応したトランザクションの完了した時点の、データを無矛盾な状態で復元することができる。

【 0 0 7 2 】

本実施形態では、ディスク 3 3 にディスク 3 1 のある時点のデータコピーが残っている場合には、ログディスク（ディスク 3 2）を参照してディスク 3 3 に対するデータ更新を行い、アプリケーションプログラム 1 6 により実行されたトランザクションの最新結果を反映することができる。

【 0 0 7 3 】

図 8 は、本実施形態において、ディスク 3 1 の内容をホスト 2 c で使えるようにするための処理のフローチャートである。

【 0 0 7 4 】

ディスク制御装置 4 3 は、ホスト 1 からディスク 3 1、3 2 が、ホスト 2 からディスク 3 4 が、プロセッサ 5 からディスク 3 2、3 3、3 4 が見えるよう記憶装置サブシステム 3 c を制御する。ディスク 3 1、3 3 の内容は、本処理の開始前のある時点において同一にしておく。

【 0 0 7 5 】

変換プログラム 6 4 は、ディスク 3 2 に格納されている更新データのうち、過去に行った処理により最後にアクセスした更新データのチェックポイント ID を保持している。変換プログラム 6 4 は、ディスク 3 2 に格納されている更新データを、保持しているチェックポイント ID に続くチェックポイント ID で特定される更新データから順に更新履歴に沿って読み出し、更新データに基づいてディスク 3 3 の内容を更新する。この更新処理は、ディスク 3 2 に記録された最新の

チェックポイント I D の更新データまで実施される。この間、アプリケーションプログラム 1 6 及びアプリケーションプログラム 2 6 は、通常の処理を継続して行うことが可能である（ステップ 1 4 0 1）。

【 0 0 7 6 】

ディスク 3 3 の更新が終了した後、変換プログラム 6 4 は、更新したディスク 3 3 の内容をディスク 3 4 に反映させるために、アプリケーション間通信機構 6 4 1 を用いてホスト 2 c のアプリケーションプログラム 2 6 に、処理の中断を要求する（ステップ 1 4 0 2）。

【 0 0 7 7 】

アプリケーションプログラム 2 6 は、プロセッサ 6 からの要求に応答して、実行中の処理を中断する（ステップ 1 1 0 1）。処理の中断後、アプリケーションプログラム 2 6 は、処理を中断したことをプロセッサ 6 に通知する（ステップ 1 5 0 1）。

【 0 0 7 8 】

変換プログラム 6 4 は、ディスク 3 3、ディスク 3 4 を duplex 状態にするために、記憶装置サブシステム 3 c に対しペア形成を指示する（ステップ 1 4 0 3）。ペア形成が終了し、ディスク 3 4 の内容がディスク 3 3 の内容と同一な duplex 状態になった後、変換プログラム 6 4 は、ペア解除を記憶装置サブシステム 3 c に対して指示し、ディスク 3 3、3 4 を simplex 状態に戻す（ステップ 1 4 0 4）。

【 0 0 7 9 】

ディスク 3 3、3 4 が simplex 状態に戻った後、変換プログラム 6 4 は、処理の完了をホスト 2 c に通知する（ステップ 1 4 0 5）。アプリケーションプログラム 2 6 は、この通知を受けると、中断していた処理を再開する（ステップ 1 1 0 3）。

【 0 0 8 0 】

以上の処理を必要とされるタイミングで繰り返すことで、アプリケーションプログラム 1 6 により使用されるディスク 3 1 の任意の時点のデータをアプリケーションプログラム 2 6 から使用することができる。

【 0 0 8 1 】

本実施形態によれば、ホスト 1 c 側で稼働中のアプリケーションプログラムの処理を中断することなくディスク 3 1 の内容をホスト 2 c 側で利用できるようにすることができる。これにより、例えば、ホスト 1 c 上で複数のデータベースプログラムが稼働している場合、あるいは、それぞれデータベースが稼働している複数のホストがあり、1 つのホストでそれらのデータを集約して処理しようとする場合などへの対応が容易となる。

【 0 0 8 2 】

例えば、ホスト 1 で複数のデータベースプログラムが稼働していて、複数のデータベース及びデータベースログが存在している場合を考える。この場合、プロセッサ 6 の変換プログラム 6 4 でこれら複数のデータベースログからデータを読み出し、プロセッサ 6 上でその内容をマージして、ディスク 3 3 に集約したデータベースを作る。ホスト 2 からは、ディスク 3 3 に集約されたデータベースを利用することが可能になる。

【 0 0 8 3 】

— 第 5 の実施形態 —

図 8 は、本発明の第 5 の実施形態における計算機システムの構成例を示すブロック図である。

【 0 0 8 4 】

本実施形態における計算機システムは、ホスト 1 d、ホスト 2 d、ホスト 1 d に接続される記憶装置サブシステム 3 0 a、及び、ホスト 2 d に接続される記憶装置サブシステム 3 0 b を含んで構成される。

【 0 0 8 5 】

ホスト 1 d、ホスト 2 d は、第 2 の実施形態におけるホスト 1 b、2 b と同様の機能を有する。ただし、ホスト 1 d とホスト 2 d は、地理的に離れた場所に設置されており、それぞれ異なる LAN 5 a、5 b に接続されている。LAN 5 a と LAN 5 b とは、例えば、インターネット 5 c により相互に接続されている。

【 0 0 8 6 】

記憶装置サブシステム 3 0 a は、ディスク 3 1 a、ホスト 1 d と接続するイン

タフェース 4 1 a、記憶装置サブシステム 3 0 b と接続するインタフェース 4 2 a、ディスク制御装置 4 3 a を有して構成される。また、記憶装置サブシステム 3 0 b は、ディスク 3 1 b、ディスク 3 2 b、ディスク 3 3 b、ホスト 2 d と接続するインタフェース 4 1 b、記憶装置サブシステム 3 0 a と接続するインタフェース 4 2 b、ディスク制御装置 4 3 b を有して構成される。記憶装置サブシステム 3 0 a と記憶装置サブシステム 3 0 b は、それぞれのインタフェース 4 2 a、4 2 b により接続される。

【 0 0 8 7 】

記憶装置サブシステム 3 0 b は、上述した各実施形態の記憶装置サブシステムと同じく、ディスクマッピング変更機能を有する。また、記憶装置サブシステム 3 0 a と記憶装置サブシステム 3 0 b は、記憶装置サブシステム 3 0 a 内のディスクの内容をインタフェース 4 2 a、4 2 b を介して記憶装置サブシステム 3 0 b 内のディスクにコピーするリモートコピー機能を備える。

【 0 0 8 8 】

リモートコピー機能は、先に説明したミラーリング機能を複数のディスク装置間で実現したものである。例として、図 8 において、ディスク 3 1 a の内容をディスク 3 2 b にコピーすることを想定する。この時、ディスク 3 1 a のことを正ディスク、ディスク 3 2 b のことを副ディスクと呼ぶ。

【 0 0 8 9 】

ユーザは記憶装置サブシステム 3 0 a に対して、ディスク 3 1 a、3 2 b をそれぞれ正／副ディスクとして関連付け、コピー開始を指示する。この指示は、ホスト 1 d からインタフェース 1 3 を介して行う。第 1 の実施形態のように記憶装置サブシステム 3 0 a、3 0 b にコンソールが接続されている場合には、ホスト 1 d からネットワーク経由でコンソールに指示を送るようにすることも可能である。

【 0 0 9 0 】

コピー開始が指示されると、ディスク制御装置 4 3 a は、インタフェース 4 2 a を介して記憶装置サブシステム 3 0 b にコピー対象となるデータを転送し、コピー処理を実施する。コピー開始時点では、ディスク 3 1 a とディスク 3 1 b の

データの内容は一致していない。ディスク 31 a からディスク 31 b へのデータコピーを行っている状態を「ペア形成中」と呼ぶ。

【0091】

ペア形成中の状態にある間、ホスト 1 d からディスク 31 a に対するリード/ライトアクセスがあると、ディスク制御装置 43 は、これを受け付けて、ディスク 30 a のアクセスを行うと共に、ホスト 1 d からのライトアクセスによりディスク 31 a の内容が更新された場合には、その更新内容をディスク 31 b に反映する。

【0092】

コピー処理が完了して、ディスク 31 a の内容がディスク 31 b に完全にコピーされた状態を「duplex状態」と呼ぶ。duplex状態の時、ホスト 1 d からディスク 31 a へライトアクセスがあると、その内容はディスク 31 b にも反映される。これらのことは、先に説明したミラーリング機能と同じである。

【0093】

duplex状態にした後、任意の時点でディスク 31 a の状態をディスク 31 b に保存しておきたい時、ユーザは、記憶装置サブシステム 3 a に「ペア解除」の指示を送る。記憶装置サブシステム 3 a は、ペア解除の指示を受けると、それ以降ディスク 31 a に対するデータの更新をディスク 31 b に反映させず、ディスク 31 b の内容をその時点の状態で保存する。この状態を「simplex状態」と呼ぶ。

【0094】

リモートコピー機能では、あるディスクに対してもう 1 台ディスクを用意して、ペアを形成してduplex状態にし、その後任意の時点でsimplex状態にすることで、その時点でのコピーを作成できる。

【0095】

ホスト 1 d のアプリケーションプログラム 16、ホスト 2 d のアプリケーションプログラム 26、及び変換プログラム 25 は、第 2 の実施の形態と同様の機能を実現する。本実施形態で、ディスク 31 a のデータをホスト 2 d で使用できるようにする時の処理は、第 2 の実施の形態と同様に、図 4 のフローチャートで示さ

れる。

【0096】

通常の処理では、ディスク31aとディスク31bがペア状態で使用される。ホスト1dは、ステップ1002、ステップ1005のペア解除、ペアの再形成では、ディスク31a、31bに対するペアの解除、再形成を行う。

【0097】

ホスト2d側では、ステップ1201において、変換プログラム25がディスク31bに保存されたデータをディスク32bにデータの形式を変換しつつコピーする。アプリケーションプログラム26は、ステップ102bのディスクマッピング変更処理で、ディスク32b、ディスク33bとのマッピングを変更して両者を入れ替え、入れ替え後のディスクを利用して処理を再開する。

【0098】

なお、本実施形態では、第2の実施形態の応用として、地理的に離れたサイトに設置されたホスト間でデータを共用する計算機システムについて説明した。同様に、第3の実施形態についても、図5のホスト1cと2cを図9に示すように地理的に離れたサイトに配置し、両者の間でデータを共用するようにすることができる。この場合、プロセッサ6は、ホスト2cと同じサイトに配置すればよい。

【0099】

以上説明した実施形態によれば、ホストやネットワークに大きな負荷を掛けることなく異なる2つのホストでデータを共用することが可能となる。

【0100】

また、ホスト上で稼働しているアプリケーションプログラムの実行を極力中断することなく、2つのホスト間で共通のファイルにアクセスすることができる。

【0101】

【発明の効果】

本発明によれば、ネットワークや計算機の負荷を極力抑えつつ2つのホスト間でデータを相互利用することができる。

【図面の簡単な説明】

【図 1】

本発明の第 1 の実施形態における計算機システムの構成例を示すブロック図である。

【図 2】

ホスト上のアプリケーションプログラムによるデータ共用のための処理の流れを示すフローチャートである。

【図 3】

本発明の第 2 の実施形態における計算機システムの構成例を示すブロック図である。

【図 4】

ホスト上のアプリケーションプログラムによるデータ共用のための処理の流れを示すフローチャートである。

【図 5】

本発明の第 3 及び第 4 の実施形態における計算機システムの構成例を示すブロック図である。

【図 6】

第 3 の実施形態において、ホスト上のアプリケーションプログラム、及びプロセッサ上の変換プログラムにより行われるデータ共用のための処理の流れを示すフローチャートである。

【図 7】

第 4 の実施形態において、ホスト上のアプリケーションプログラム、及びプロセッサ上の変換プログラムにより行われるデータ共用のための処理の流れを示すフローチャートである。

【図 8】

本発明の第 5 の実施形態における計算機システムの構成例を示すブロック図である。

【図 9】

本発明の第 3 の実施形態における計算機システムの応用構成例を示すブロック

図である。

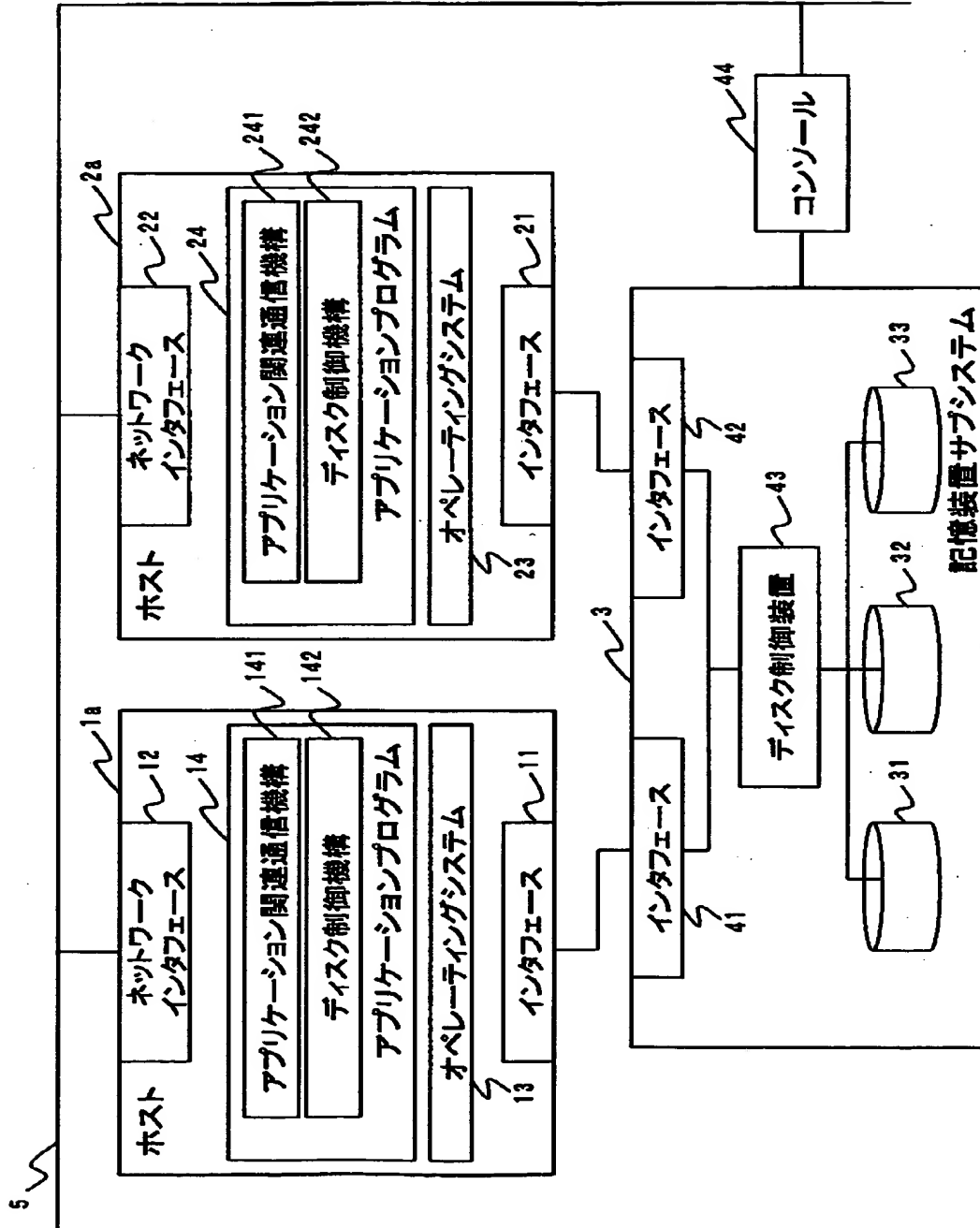
【符号の説明】

1 a、1 b、1 c、1 d、2 a、2 b、2 c、2 d…ホストコンピュータ、
3、3 a、3 b…記憶装置サブシステム、
5…ネットワーク、
1 4、1 6、2 4、2 6…アプリケーションプログラム、
2 5、6 4…変換プログラム、
3 1、3 2、3 3、3 4…ディスク、
4 3…ディスク制御装置、
4 4…コンソール。

【書類名】 図面

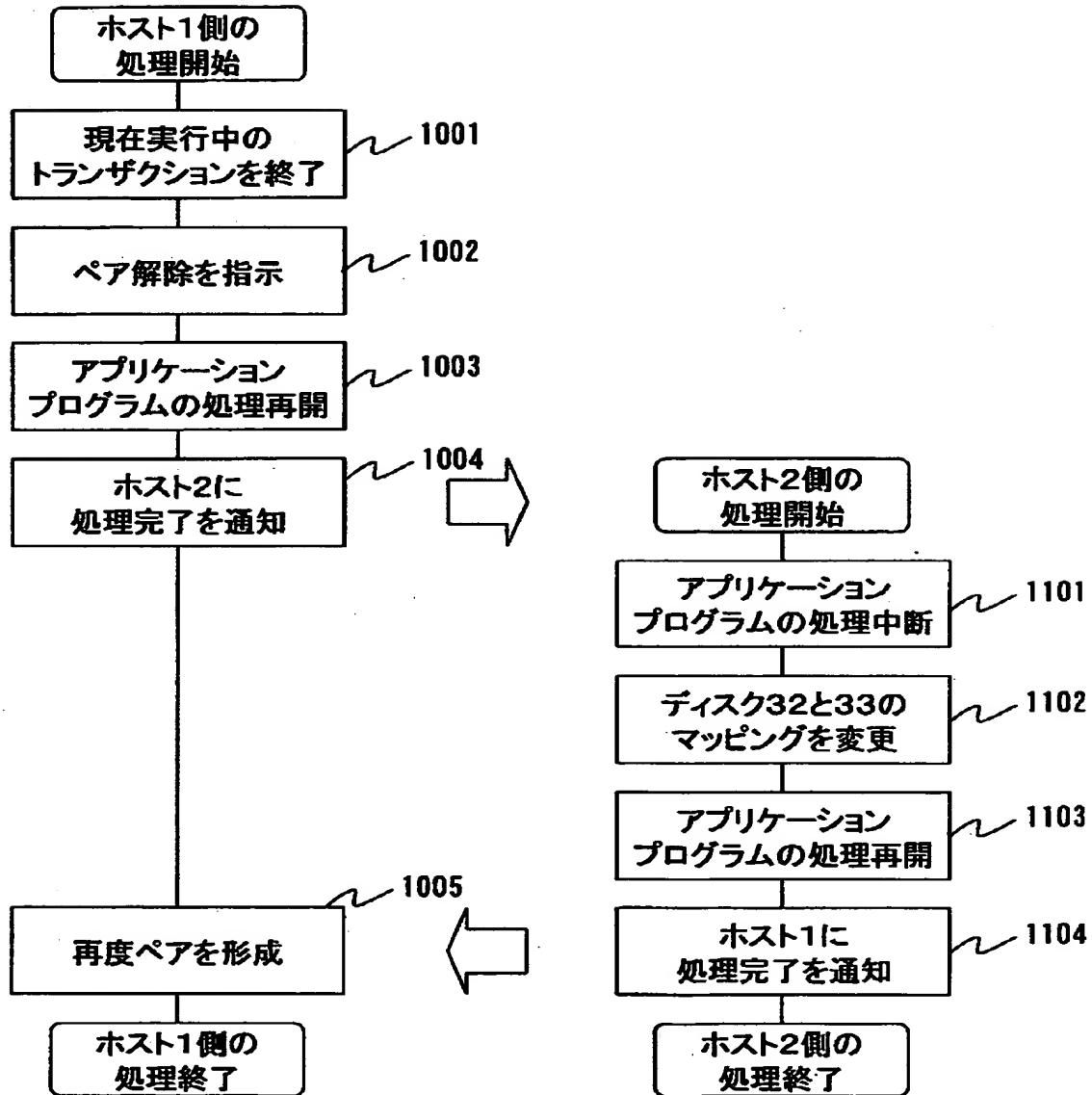
【図 1】

図 1

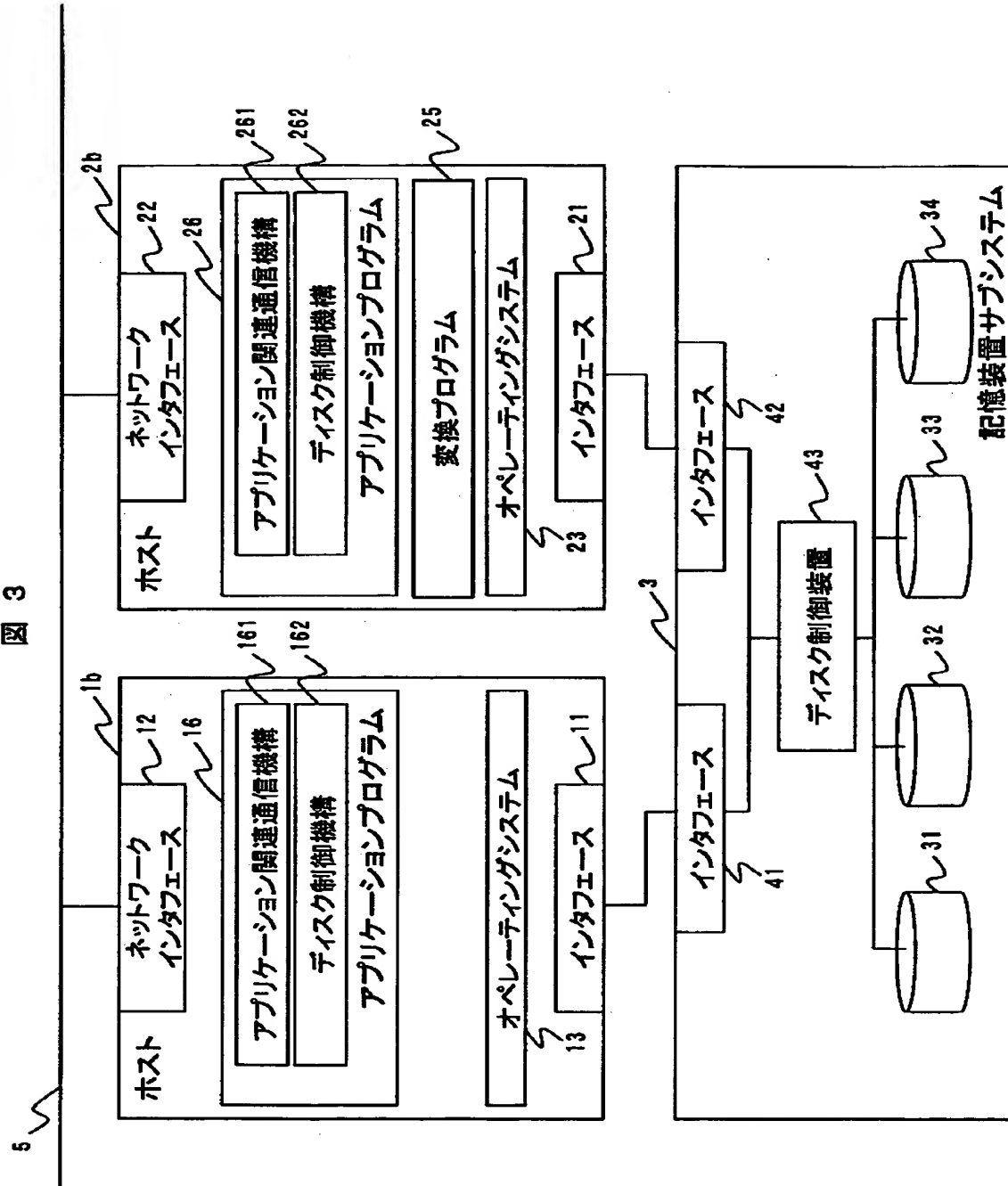


【図 2】

図 2

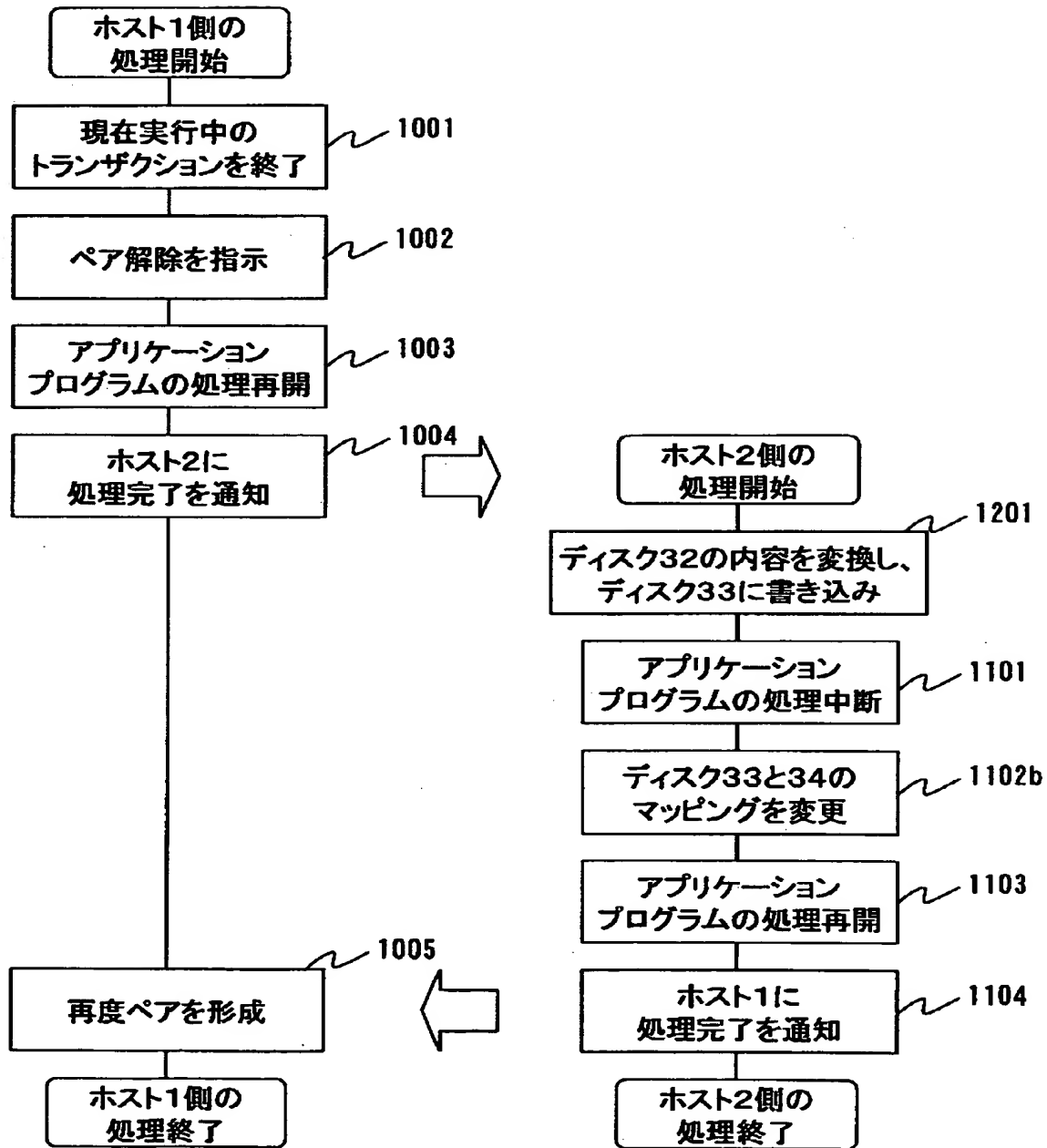


【図 3】



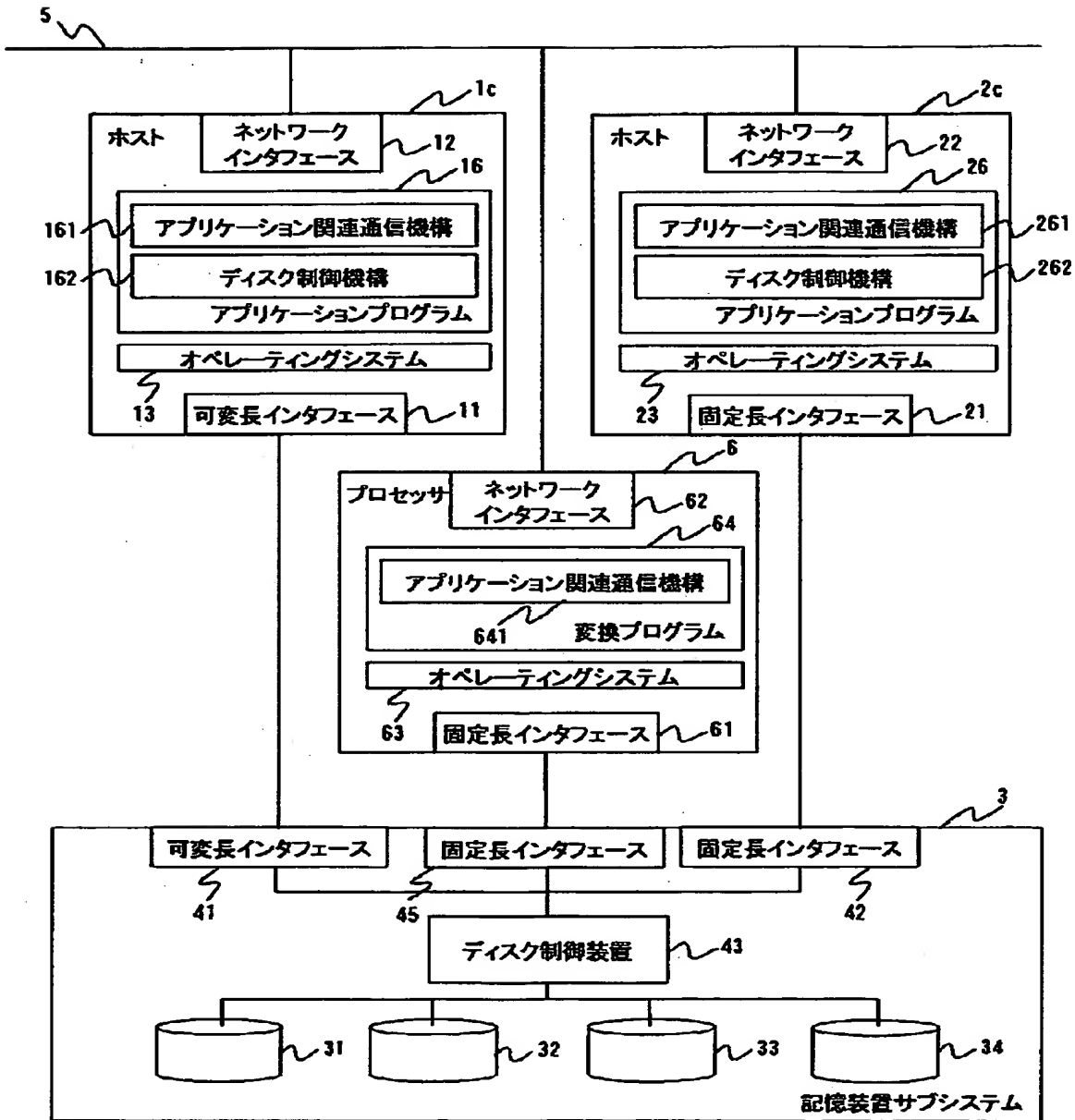
【図 4】

図 4



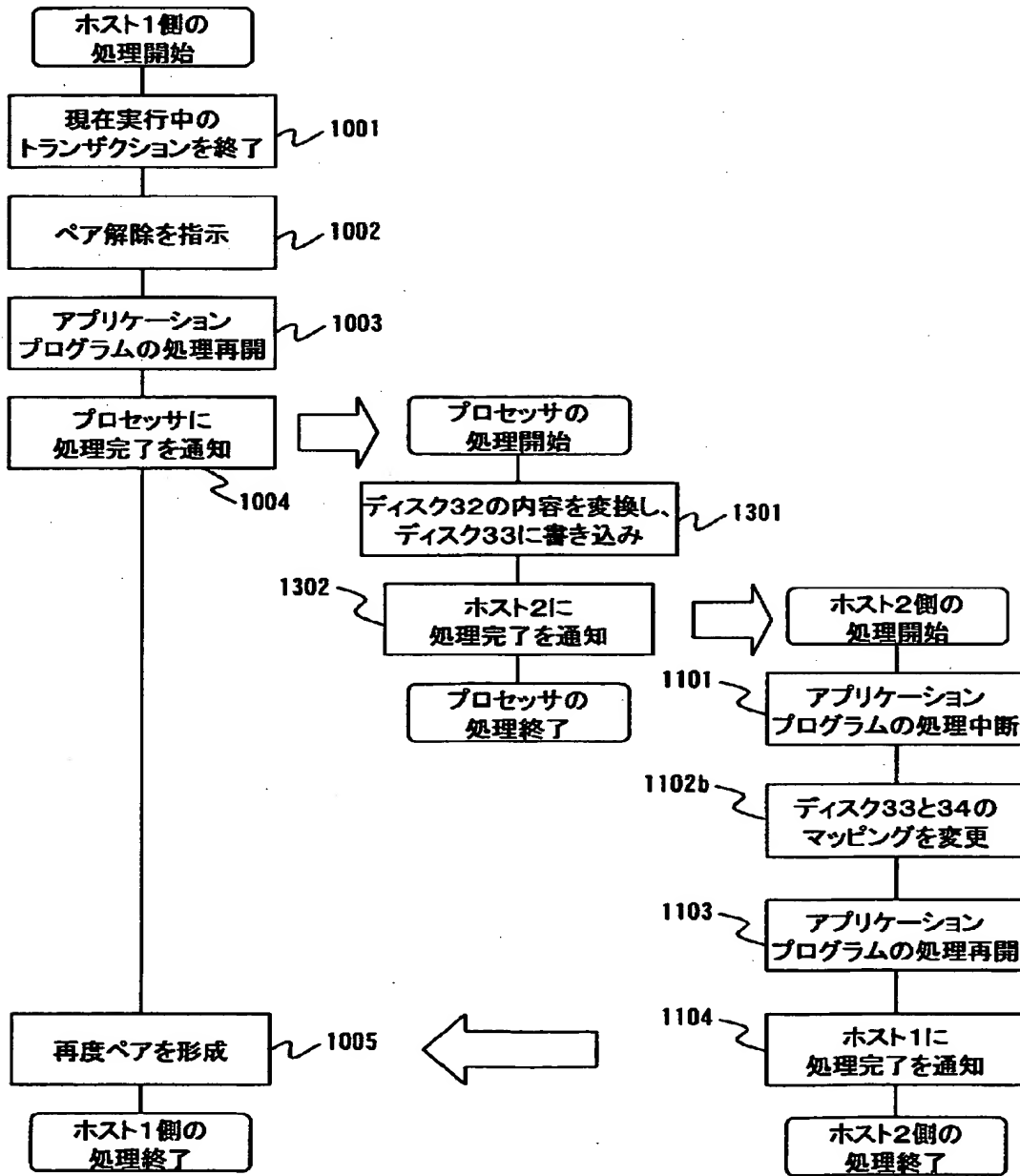
【図 5】

図 5



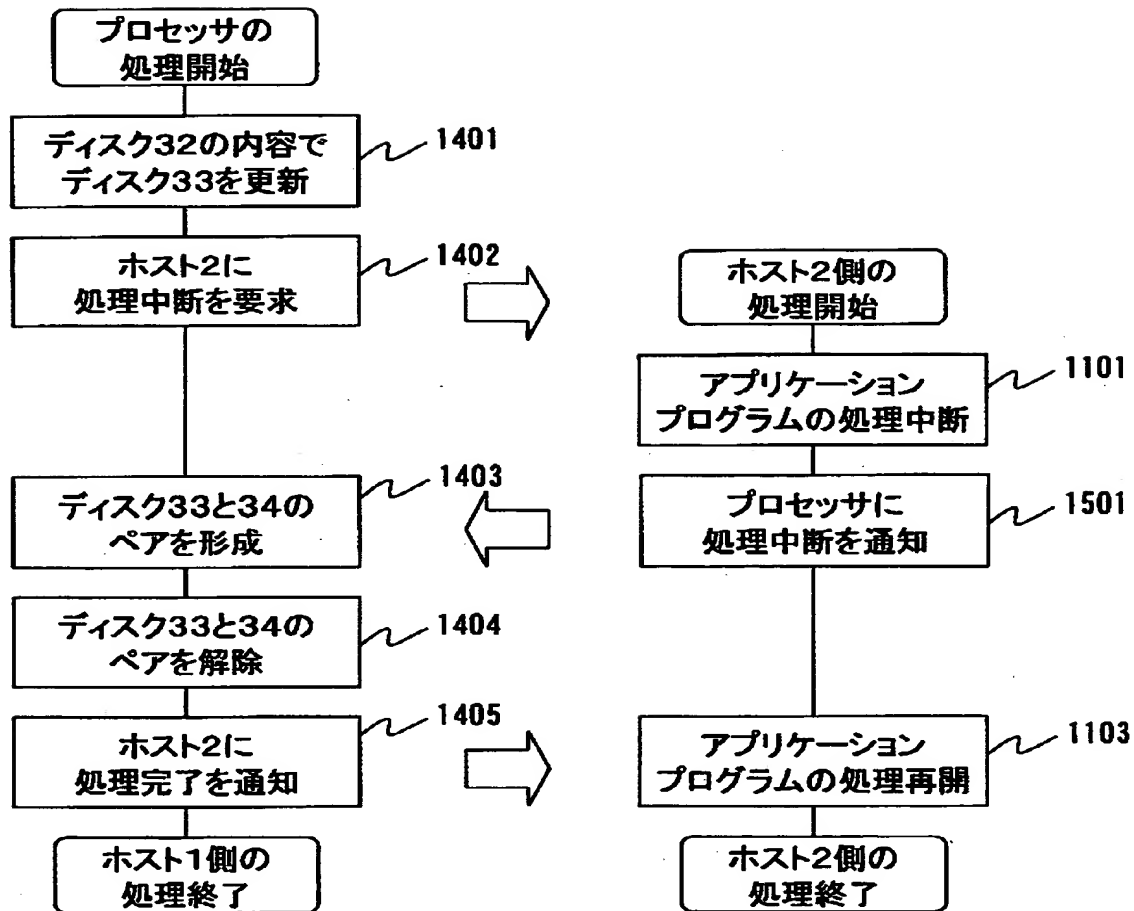
【図 6】

図 6



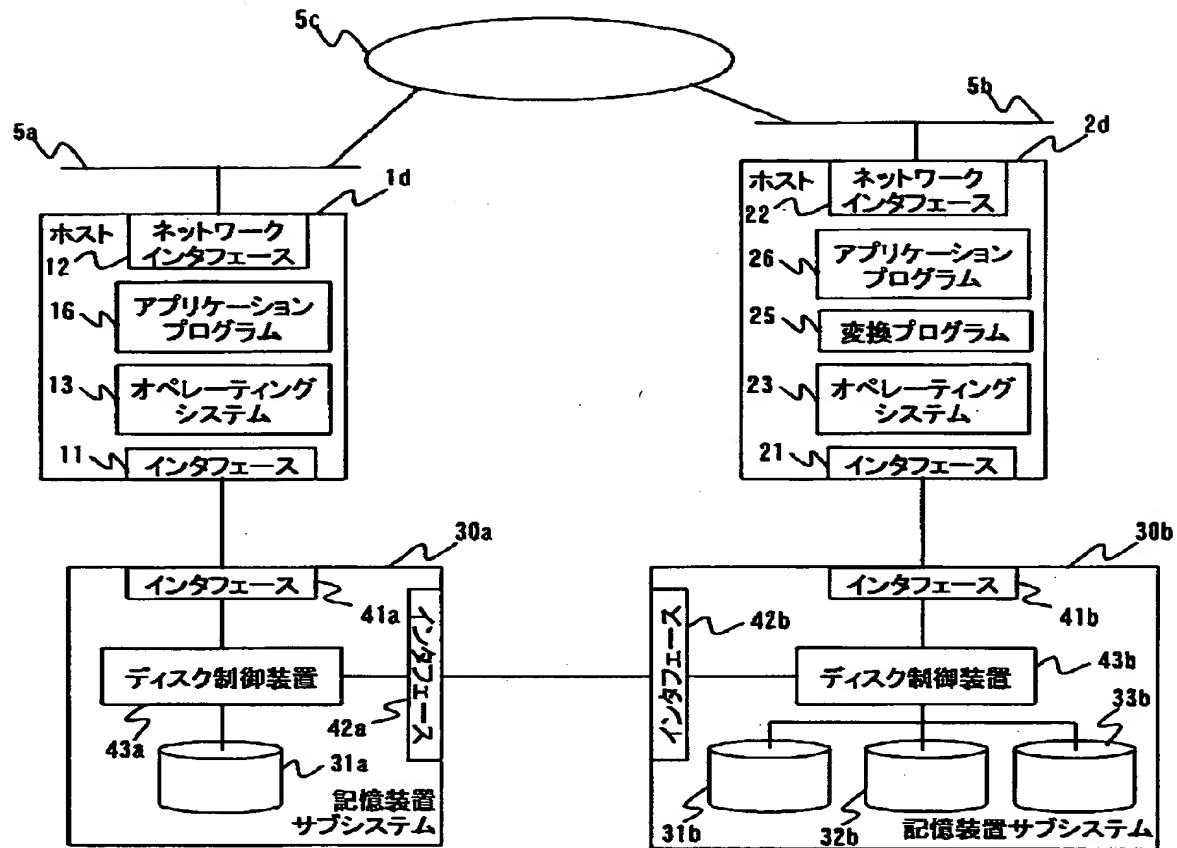
【図 7】

図 7



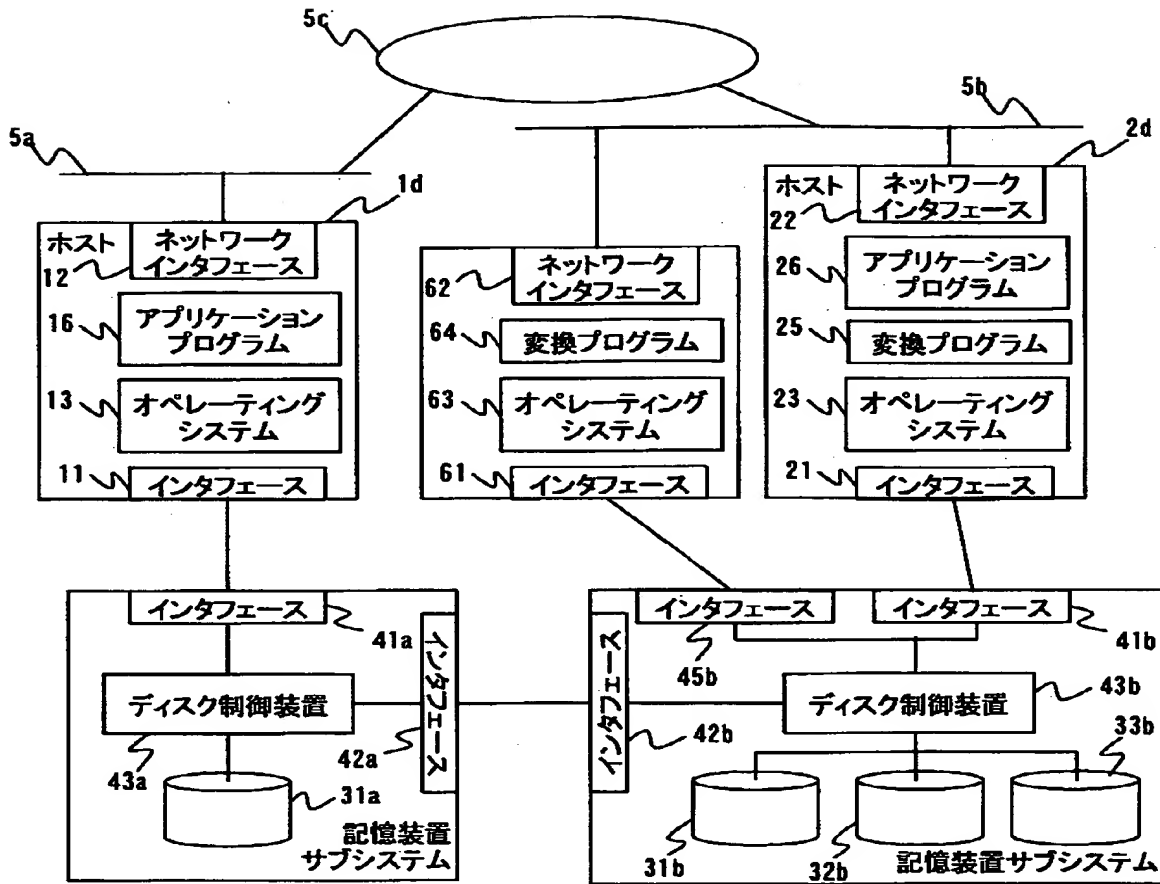
【図8】

図 8



【図 9】

図 9



【書類名】 要約書

【要約】

【課題】

複数の計算機間でのデータの共有を計算機やネットワークへの負荷を押さえつつ実現する。

【解決手段】

記憶装置サブシステム 3 は、アプリケーションプログラム 1 4 の通常処理によるディスク 3 1 のデータの更新に対し、その更新内容をディスク 3 2 にコピーする。ディスク 3 1 の内容をホスト 2 a から利用できるようにする際、アプリケーションプログラム 1 4 は、ディスク 3 2 へのコピー処理を中断させる。ホスト 2 a のアプリケーションプログラム 2 4 は、ディスク 3 2 とディスク 3 3 を交換して、ディスク 3 3 に替えてディスク 3 2 を利用できるようにする。その後、アプリケーションプログラム 2 4 は、入れ替えられたディスクを用いて処理を行う。

【選択図】 図 1

出 願 人 履 歴 情 報

識別番号 [000005108]

1. 変更年月日 1990年 8月31日
[変更理由] 新規登録
住 所 東京都千代田区神田駿河台4丁目6番地
氏 名 株式会社日立製作所